

日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

48864-040  
June 12, 2001  
KAMIMURA, ET AL  
McDermott, Will & Emery

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出願年月日  
Date of Application:

2000年 6月13日

出願番号  
Application Number:

特願2000-176678

出願人  
Applicant(s):

ミノルタ株式会社

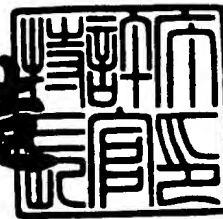


CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 3月 2日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出訴番号 出証特2001-3014273

【書類名】 特許願

【整理番号】 TL03699

【提出日】 平成12年 6月13日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/14

【発明の名称】 アニメーション通信方法およびシステム並びにそれに用  
いる端末装置

【請求項の数】 9

【発明者】

【住所又は居所】 大阪府大阪市中央区安土町二丁目3番13号 大阪国際  
ビル ミノルタ株式会社内

【氏名】 上村 時男

【発明者】

【住所又は居所】 大阪府大阪市中央区安土町二丁目3番13号 大阪国際  
ビル ミノルタ株式会社内

【氏名】 田中 譲

【特許出願人】

【識別番号】 000006079

【氏名又は名称】 ミノルタ株式会社

【代理人】

【識別番号】 100086933

【弁理士】

【氏名又は名称】 久保 幸雄

【電話番号】 06-6304-1590

【手数料の表示】

【予納台帳番号】 010995

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9716123

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 アニメーション通信方法およびシステム並びにそれに用いる端末装置

【特許請求の範囲】

【請求項 1】

相手のアニメーションを見ながら会話を行うためのアニメーション通信方法であって、

ネットワークを介して接続された相手の端末装置から送信される音声に基づいて、画像データを動作させるための動作制御データを生成するステップと、

受信した音声を再生するとともに、前記動作制御データに基づいて画像データを動作させて生成される 3 次元のアニメーションを表示するステップと、

を有してなることを特徴とするアニメーション通信方法。

【請求項 2】

相手のアニメーションを見ながら会話を行うためのアニメーション通信システムに用いられる端末装置であって、

ネットワークを介して接続された相手の端末装置から送信される音声を受信する手段と、

受信した音声に基づいて、画像データを動作させるための動作制御データを生成する手段と、

受信した音声を再生するとともに、前記動作制御データに基づいて画像データを動作させて生成される 3 次元のアニメーションを表示する手段と、

を有してなることを特徴とする端末装置。

【請求項 3】

相手のアニメーションを見ながら会話を行うためのアニメーション通信システムであって、

ホストコンピュータと複数の端末装置とを有し、

前記端末装置には、音声を送受信する手段と、3 次元の画像データを受信する手段と、前記画像データを動作させるための動作制御データを受信する手段と、前記動作制御データに基づいて前記画像データを動作させて生成される 3 次元の

アニメーションを表示する手段とが設けられ、

前記ホストコンピュータには、音声を送受信する手段と、受信した音声に基づいて前記動作制御データを生成する手段と、通信の一方の端末装置に係る前記画像データおよび前記動作制御データを通信の他方の端末装置に送信する手段とが設けられ、

てなることを特徴とするアニメーション通信システム。

【請求項 4】

相手のアニメーションを見ながら会話を行うためのアニメーション通信システムであって、

ホストコンピュータと複数の端末装置とを有し、

前記端末装置には、音声を送受信する手段と、画像データを受信する手段と、前記画像データを動作させるための動作制御データを受信する手段と、前記動作制御データに基づいて前記画像データを動作させて生成されるアニメーションを表示する手段とが設けられ、

前記ホストコンピュータには、音声を受信する手段と、受信された音声を他の自然言語に翻訳する手段と、翻訳された音声を送信する手段と、翻訳された音声に基づいて前記動作制御データを生成する手段と、通信の一方の端末装置に係る前記画像データおよび前記動作制御データを通信の他方の端末装置に送信する手段とが設けられ、

てなることを特徴とするアニメーション通信システム。

【請求項 5】

相手のアニメーションを見ながら会話を行うためのアニメーション通信システムに用いられる端末装置であって、

音声を送受信する手段と、

3 次元の画像データを受信する手段と、

前記画像データを動作させるための動作制御データを受信する手段と、

前記動作制御データに基づいて前記画像データを動作させて生成されるアニメーションを表示する手段と、

が設けられてなることを特徴とする端末装置。

【請求項 6】

相手のアニメーションを見ながら会話を行うためのアニメーション通信システムに用いられるホストコンピュータであって、

音声を送受信する手段と、

受信された音声を他の自然言語に翻訳する手段と、

翻訳された音声を送信する手段と、

翻訳された音声に基づいて、画像データを動作させるための動作制御データを生成する手段と、

通信の一方の端末装置に係る前記画像データおよび前記動作制御データを通信の他方の端末装置に送信する手段と、

が設けられてなることを特徴とするホストコンピュータ。

【請求項 7】

相手のアニメーションを見ながら会話を行うためのアニメーション通信システムであって、

ホストコンピュータと複数の端末装置とを有し、

前記端末装置には、音声を送受信する手段と、画像データを記憶する記憶手段と、画像データを送受信する手段と、受信した音声に基づいて受信した画像データを動作させるための動作制御データを生成する手段と、前記動作制御データに基づいて受信した画像データを動作させて生成されるアニメーションを表示する手段とが設けられ、

前記ホストコンピュータには、音声を受信する手段と、受信された音声を他の自然言語に翻訳する手段と、翻訳された音声を送信する手段とが設けられ、

てなることを特徴とするアニメーション通信システム。

【請求項 8】

相手のアニメーションを見ながら会話を行うためのアニメーション通信システムに用いられる端末装置であって、

音声を送受信する手段と、

3 次元の画像データを記憶する記憶手段と、

画像データを送受信する手段と、

受信した音声に基づいて受信した画像データを動作させるための動作制御データを生成する手段と、

生成した動作制御データに基づいて受信した画像データを動作させて生成されるアニメーションを表示する手段と、

が設けられてなることを特徴とする端末装置。

【請求項 9】

前記画像データは、動作可能に構造化された 3 次元モデルであり、

前記動作制御データに基づいて 3 次元モデルの構造化された部分を動作させる

請求項 2 ないし 8 のいずれかに記載のシステム、ホストコンピュータ、または端末装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、相手のアニメーションを見ながら会話を行うためのアニメーション通信方法およびシステム、並びにそれに用いる端末装置およびホストコンピュータに関する。

【0002】

【従来の技術】

従来より、テレビ電話またはテレビ会議システムなどのように、各端末装置間において音声と画像を送受信することにより、相手の顔を見ながら会話を行うことのできる通信システムが実用化されている。

【0003】

しかし、画像はデータ量が多いため、画像の送受信を行うには通信容量の大きな通信回線が必要である。そのため、一般の電話回線を介して画像の送受信を行った場合には、毎秒数コマ程度の画像の送受信しか行うことができず、満足のいく動画を表示することができない。高速の専用線を介した場合には、自然な動きに近い動画を表示することが可能であるが、通信コストが高いため一般の家庭に普及するには至っていない。

## 【0004】

そこで、通信量を軽減するために、顔の全体および一部分の画像を低解像度で予め作成してデータベースに登録しておき、会話の開始時に顔の全体の画像を相手の端末装置の表示面に表示し、表情の変化のあった部分に対応する顔の一部分の画像のみをデータベースから端末装置へダウンロードして表示する方法が提案されている（特開平10-200882号）。

## 【0005】

## 【発明が解決しようとする課題】

しかし、上に述べた従来の方法では、通信量の軽減を図ることはできるが、画像が低解像度であることと、複数の2次元画像を連続的に組み合わせて表示していることから、人の表情などの自然な動きを表現することが困難である。

## 【0006】

また、通信システムを用いて会話を行う者同士が共通の言語を理解できる必要があるので、言語の異なる者同士では上のような通信システムを用いることができない。

## 【0007】

本発明は、通信量を軽減させ、且つ自然に近いスムーズな相手のアニメーションを見ながら会話を行うことを可能にすることを目的とする。

また、言語の異なる者同士であっても、自然に近いスムーズな相手のアニメーションを見ながら会話を行うことを可能にすることを他の目的とする。

## 【0008】

## 【課題を解決するための手段】

本発明に係る方法は、相手のアニメーションを見ながら会話を行うためのアニメーション通信方法であって、ネットワークを介して接続された相手の端末装置から送信される音声に基づいて、画像データを動作させるための動作制御データを生成するステップと、受信した音声を再生するとともに、前記動作制御データに基づいて画像データを動作させて生成される3次元のアニメーションを表示するステップと、を有してなる。

## 【0009】



本発明に係るシステムは、ホストコンピュータと複数の端末装置とを有し、前記端末装置には、音声を送受信する手段と、3次元の画像データを受信する手段と、前記画像データを動作させるための動作制御データを受信する手段と、前記動作制御データに基づいて前記画像データを動作させて生成される3次元のアニメーションを表示する手段とが設けられ、前記ホストコンピュータには、音声を送受信する手段と、受信した音声に基づいて前記動作制御データを生成する手段と、通信の一方の端末装置に係る前記画像データおよび前記動作制御データを通信の他方の端末装置に送信する手段とが設けられてなる。

#### 【0010】

本発明の他の形態によると、ホストコンピュータには、受信された音声を他の自然言語に翻訳する手段と、翻訳された音声を送信する手段と、翻訳された音声に基づいて前記動作制御データを生成する手段とが設けられる。

#### 【0011】

本発明の他の形態によると、端末装置には、画像データを記憶した記憶手段が設けられ、その記憶手段から選択された画像データが通信の相手の端末装置に送信される。

#### 【0012】

本発明において、端末装置として、電話機、携帯電話機、モバイル端末、テレビ電話端末、パーソナルコンピュータなど、画像の表示と通信を行うことのできる種々の装置を用いることができる。

#### 【0013】

#### 【発明の実施の形態】

以下に示す2つの実施形態では、通信の相手の実物の顔画像の代わりに、相手のアバター（身代わり）である顔のアニメーションを見ながら会話を行う通信システム1を例として説明する。通信システム1では、端末装置としてパーソナルコンピュータを用いているが、電話機、携帯電話機、モバイル機器など、他の種々の通信機器を用いることが可能である。

#### 〔第一の実施形態〕

図1は本発明に係る第一の実施形態の通信システム1の全体の構成を示すプロ

ック図、図2は第一の実施形態における端末装置2の側に記憶されているプログラムおよびデータの例を示す図、図3は第一の実施形態におけるホストコンピュータ3の側に記憶されているプログラムおよびデータの例を示す図、図4はアバターとなる顔画像データFGDの例を示す図である。

## 【0014】

図1に示すように、通信システム1は、端末装置2、ホストコンピュータ3、およびネットワーク4などから構成される。端末装置2は複数設けられており、図1においてはそのうちの2つの端末装置2A、2Bが示されている。

## 【0015】

端末装置2は、処理装置21、表示装置22a、スピーカ22b、マウス23a、キーボード23b、マイク23c、通信制御装置24、CD-ROM装置25、フロッピーディスク装置26、および磁気ディスク装置27などから構成される。

## 【0016】

処理装置21は、CPU21a、RAM21b、およびROM21cなどから構成され、端末装置2における一連の処理を実行する。

RAM21bにはプログラムまたはデータなどが一時的に記憶され、ROM21cには、プログラム、および端末装置2のハードウェアの設定情報などが記憶される。CPU21aは、それらのプログラムを実行する。

## 【0017】

表示装置22aは、顔のアニメーションなどを表示する。スピーカ22bは、会話の相手の音声を出力する。マウス23aおよびキーボード23bは、端末装置2の操作などに用いられる。マイク23は音声を入力する。

## 【0018】

通信制御装置24は、顔の3次元データである顔画像データFGD、音声が出力されるタイミングに合わせて顔画像データFGDが動作するように制御する動作制御データDSD、音声をデジタル変換した音声データSND、その他のデータの送受信を制御する。CD-ROM装置25、フロッピーディスク装置26、および磁気ディスク装置27は、データおよびプログラムを記憶する。

## 【 0 0 1 9 】

ホストコンピュータ 3 は、処理装置 3 1、表示装置 3 2、マウス 3 3 a、キーボード 3 3 b、通信制御装置 3 4、CD-ROM 装置 3 5、フロッピーディスク装置 3 6、および磁気ディスク装置 3 7 などから構成される。

## 【 0 0 2 0 】

処理装置 3 1 は、CPU 3 1 a、RAM 3 1 b、および ROM 3 1 c などから構成される。処理装置 3 1 それ自体の構成および機能は上に述べた処理装置 2 1 と同様である。

## 【 0 0 2 1 】

ネットワーク 4 は、公衆回線、専用線、LAN、無線回線、またはインターネットなどによって構成される。端末装置 2 とホストコンピュータ 3 は、ネットワーク 4 を介して接続される。

## 【 0 0 2 2 】

図 2 に示すように、端末装置 2 の磁気ディスク装置 2 7 には、端末装置の基本プログラムである OS 2 s、通信システム 1 の端末装置側のアプリケーションプログラムである端末通信プログラム 2 p、その他必要なプログラムおよびデータが記憶されている。

## 【 0 0 2 3 】

端末通信プログラム 2 p は、基本処理プログラム 2 p k、表示処理プログラム 2 p h などのプログラムまたはモジュールを含んでいる。基本処理プログラム 2 p k は、OS 2 s との連携、顔画像データ FGD の選択などのユーザの操作に関する処理などを行う。表示処理プログラム 2 p h は、動作制御データ DSD に基づいて顔画像データ FGD を動作させ、アニメーションを生成する。

## 【 0 0 2 4 】

これらのプログラムは、適宜、RAM 2 1 b にロードされ、CPU 2 1 a によって実行される。また、受信した顔画像データ FGD、動作制御データ DSD、および音声データ SND などのデータは、RAM 2 1 b に記憶される。また、必要に応じて磁気ディスク装置 2 7 に記憶される。

## 【 0 0 2 5 】

このように、種々のプログラムがRAM21b上で実行されることによって、アニメーションの表示を行うための一連のシステムである表示制御部EM1が実現される。

## 【0026】

図3に示すように、ホストコンピュータ3の磁気ディスク装置37には、ホストコンピュータ3の基本プログラムであるOS3s、通信システム1のホストコンピュータ側のアプリケーションプログラムであるホスト通信プログラム3p、その他必要なプログラムおよびデータが記憶されている。また、顔画像データFGDを蓄積する顔画像データベースFDBが設けられている。

## 【0027】

ホスト通信プログラム3pは、基本処理プログラム3pk、動作制御プログラム3pd、翻訳プログラム3pyなどのプログラムまたはモジュールを含んでいる。基本処理プログラム3pkは、OS3sとの連携、アニメーションエンジンEM2、および翻訳エンジンEM3などを統括して制御する。動作制御プログラム3pは、音声データSNDに基づいて動作制御データDSDを生成する。動作制御データDSDとは、音声データSNDに基づく音声が出力されるタイミングに合わせて、顔画像データFGDが動作するように制御するための制御情報である。翻訳プログラム3pyは、ある自然言語の音声データSNDを他の自然言語の音声データSNDに翻訳する。

## 【0028】

これらのプログラムは、適宜、RAM31bにロードされ、CPU31aによって実行される。また、受信した音声データSNDなどのデータはRAM31bに記憶される。

## 【0029】

このように、種々のプログラムがRAM31b上で実行されることによって、動作制御データDSDを生成する一連のシステムであるアニメーションエンジンEM2、および、音声データSNDを他の言語に翻訳する一連のシステムである翻訳エンジンEM3が実現される。

## 【0030】

なお、翻訳元の音声データを「音声データ S N D 1」とし、翻訳後の音声データを「音声データ S N D 2」として、これらを区別することがある。言語の自動翻訳に関しては、例えば特開平 1 - 2 1 1 7 9 9 号を参照することができる。

## 【 0 0 3 1 】

図 4 に示すように、顔画像データ F G D は、人の頭部などを、口、目、鼻、耳などの構成部分、および皮膚、筋肉、骨格などが動作可能なように構造化された 3 次元モデルによって表現したデータである。図 4 においては、皮膚を表現する多数のポリゴンと、顔の動きを制御するための複数の制御点 P N T が示されている。顔画像データ F G D は、図 4 で示される皮膚のモデルと、図示しない筋肉および骨格のモデルとによって構造化されている。

## 【 0 0 3 2 】

顔画像データ F G D として、自分の顔の他、芸能人またはスポーツ選手などの有名人、歴史上の人物、動物、漫画の登場人物（キャラクタ）など、実在または架空の種々のものを用いることができる。

## 【 0 0 3 3 】

このような顔画像データ F G D を作成する方法として、2 次元で撮影された通常の顔写真に基づいて構造化顔画像を推定する方法、3 次元計測装置を用いて人物または像を 3 次元測定し、得られた 3 次元データに基づいて顔画像データ F G D を作成する方法など、種々の公知の方法を用いることができる。

## 【 0 0 3 4 】

なお、これらの方法の例が、特開平 8 - 2 9 7 7 5 1 号、特開平 1 1 - 3 2 8 4 4 0 号、および本出願人が提案する特願 2 0 0 0 - 9 0 6 2 9 号などに示されている。

## 【 0 0 3 5 】

次に、一方の端末装置 2 A のユーザと他方の端末装置 2 B のユーザとの会話に際し、通信システム 1 で行われる処理および操作について、フローチャートを参照して説明する。

## 【 0 0 3 6 】

図 5 は第一の実施形態における通信システム 1 の処理の流れを示すフローチャ

ート、図6は端末装置2の処理の流れを示すフローチャート、図7はホストコンピュータ3の処理の流れを示すフローチャートである。

【0037】

まず、端末装置2Aと2Bとの間で通信を確立する(#10)。通信を確立するに当たり、例えば、一方の端末装置2Aからホストコンピュータ3に対し、他方の端末装置2Bへの接続要求を送る。ホストコンピュータ3は、端末装置2Aから接続要求のあったことを端末装置2Bに通知する。端末装置2Bは、接続を許可する場合に、それに応じた通知を行う。その他、公知の種々のプロトコルを用いることができる。

【0038】

通信が確立すると、図5に示すように、ホストコンピュータ3から各端末装置2に対し、それぞれの相手の顔画像データFGDを送信する(#11)。つまり、端末装置2Aに対しては、端末装置2Bのユーザの選択した顔画像データFGDを、端末装置2Bに対しては、端末装置2Aのユーザの選択した顔画像データFGDを、それぞれ送信する。各ユーザは、顔画像データベースFDBの中から、または予め各ユーザ用に登録したもののの中から、自分の好みの顔画像データFGDを選択する。そのような選択に際し、選択可能な顔画像データFGDの一覧表を各ユーザの表示面に表示させるようにしてもよいし、ユーザが番号などで指定するようにしてもよい。また、各ユーザが予め指定した1つの顔画像データFGDが送信されるようにしてもよい。

【0039】

各ユーザは会話を開始する(#12)。会話が行われると、一方の端末装置2から他方の端末装置2に向けて音声データSNDが送信される。

このとき、ユーザは、自分の発する言語および聞き取りに使う言語をホストコンピュータ3に対して指定することができる。例えば、英語で会話を行いたい場合には、話し言語および聞き取り言語とも英語を指定する。日本語で話し、英語で音声を聞くように指定することもできる。会話の途中で言語の指定を変更することも可能である。

【0040】

ホストコンピュータ3は、端末装置2からの言語の指定に応じて、これらの間の会話において翻訳を行う必要があるか否かを判断する（#13）。一方の話す言語と他方の聞く言語とが異なっている場合に、翻訳が必要であると判断する。なお、ユーザからの言語の指定がない場合には、特定の言語、例えば日本語であると判断する。

## 【0041】

翻訳が必要な場合には、ホストコンピュータ3は、翻訳エンジンEM3によって翻訳を行う（#14）。翻訳によって、音声データSND1から音声データSND2が生成される。

## 【0042】

音声データSNDに基づいて、動作制御データDSDが生成される（#15）。翻訳が行われた場合は、翻訳後の音声データSND2に基づいて動作制御データDSDが生成される。

## 【0043】

動作制御データDSDの生成に当たって、例えば、音声データSNDから音素などの情報を抽出し、単語または感情などを求め、顔画像データFGDの各制御点PNTの動きを計算して動作制御データDSDを生成する。

## 【0044】

なお、ユーザの感情を受信した音声データSNDから抽出するのではなく、ユーザが端末装置のキーボード23bなどを操作して直接に指定してもよい。その場合には、端末装置2からは、例えば「微笑む」「怒る」などの感情を表す制御データを送信することとなる。このようにすると、ユーザが疲れているときでも、会話の相手には元気な様子のアニメーションを表示することができる。

## 【0045】

ホストコンピュータ3から端末装置2に対し、音声データSNDおよび動作制御データDSDを送信する（#16）。

端末装置2では、受信した音声データSNDをスピーカ22bから出力するとともに、受信した動作制御データDSDに基づいて、最初に受信した顔画像データFGDを動作させ、これによってアニメーションを生成し、生成したアニメー

ションを表示装置 22a に表示する。( # 17 )。

【 0046 】

例えば、端末装置 2A のユーザが「おはよう」と言うと、ホストコンピュータ 3 は顔画像データ FGD に「おはよう」の口の動きをさせるための動作制御データ DSD を生成して端末装置 2B に送信する。端末装置 2B では、端末装置 2A のユーザの「おはよう」の声がスピーカ 22b から聞こえ、表示装置 22a には「おはよう」の声に合わせて端末装置 2A のユーザの顔画像データ FGD が「おはよう」の形に開閉する。

【 0047 】

また、ホストコンピュータ 3 は、「おはよう」のトーンなどから端末装置 2A のユーザの感情を分析する。そこで、例えば端末装置 2A のユーザがにこやかな雰囲気であることが分かったと、顔画像データ FGD の目および顔全体を動かして微笑ませる動作制御データ DSD を生成し、端末装置 2B に送信する。端末装置 2B の表示装置 22a には、端末装置 2A のユーザが微笑みながら「おはよう」と言っているアニメーションが表示される。

【 0048 】

このように、各ユーザは、互いに相手の話した音声聞き、且つ相手の話に基づいて表情の変化するアニメーションを見ることができる。

ユーザが通信の切断を要求するまで ( # 18 )、上に述べた処理が繰り返され、ユーザ間においてアニメーションを見ながらの会話が行われる。

【 0049 】

図 6 に示すように、各端末装置 2 においては、ホストコンピュータ 3 から相手のユーザの顔画像データ FGD を受信する ( # 21 )。各ユーザが話をする ( # 22 でイエス)、音声データ SND をホストコンピュータ 3 へ送信する ( # 23 )。動作制御データ DSD および音声データ SND を受信すると ( # 24 でイエス)、それらによる音声を出力しアニメーションを表示する ( # 25 )。

【 0050 】

図 7 に示すように、ホストコンピュータ 3 においては、各端末装置 2 に相手のユーザの顔画像データ FGD を送信する ( # 31 )。端末装置 2 から音声データ



SNDを受信すると（＃32）、翻訳の必要がある場合に翻訳を行った後（＃33、34）、動作制御データDSDを生成し（＃35）、動作制御データDSDおよび音声データSNDを相手側の端末装置2へ送信する（＃36）。

【0051】

なお、3つ以上の端末装置2の間で通信を行うことも可能である。その場合には、各ユーザに対して、他の全てのユーザの顔画像データFGDを送信する。各ユーザの音声は、それに基づく動作制御データDSDとともに、他の全てのユーザの端末装置2に送信される。各端末装置2では、受信した複数の顔画像データFGDに基づくアニメーションの中から、話をしているユーザに対応するアニメーションのみを選択して表示すればよい。また、全てのユーザのアニメーションを同時に表示するようにしてもよい。それらを切り換えるようにしてもよい。

【0052】

第一の実施形態の通信システム1によると、データ量の多い顔画像データFGDは最初に1回のみ送信され、その後は動作制御データDSDのみが送信されるので、通信量が軽減するとともに、自然に近いスムーズな相手のアニメーションを見ながら会話を行うことが可能となる。

【0053】

顔画像データFGDが構造化された3次元モデルによって表現されたデータであり、表示面には3次元のアニメーションが表示されるので、実物の画像に近いリアルな画像が表示される。

【0054】

ホストコンピュータ3にアニメーションエンジンEM2が設けられているので、端末装置2における処理の負荷が軽減される。

また、ホストコンピュータ3による翻訳サービスを仲介させることで、他国語を話す相手との間でも違和感のない会話を行うことができる。

【0055】

特に、翻訳後の音声データSNDに基づいて動作制御データDSDを生成するので、翻訳された音声とアニメーションの動きとを旨く一致させることができる。

## 【 0 0 5 6 】

例えば、言語の違いによって唇や顔の動きが異なるので、実物の顔画像を表示した場合には音声は翻訳されても顔の動きはそのままであるが、本実施形態によれば、翻訳後の音声に唇や顔の動きを合わせることができるので、翻訳された音声とともに表情を正確に表現した自然なアニメーションを会話の相手に表示することができる。

## 【 0 0 5 7 】

また、外国映画の吹替えに見られるように、異なる言語間の画像の動きと音声の響きの不一致からくる不自然さ、または音声の長さの違いからくる不自然さが解消される。

## 【 0 0 5 8 】

第一の実施形態において、翻訳後の音声データ S N D 2 とともに、翻訳前の音声データ S N D 1 を送信してもよい。これにより、音声の多重化を図ることができる。原文を確認しながら翻訳された音声を聞くことができる。

## 【 0 0 5 9 】

また、翻訳後の音声データ S N D 2 とともに、そのテキストデータを送信してもよい。これにより、端末装置 2 において、アニメーションの表示とともに、翻訳文を表示することができる。

## 【 0 0 6 0 】

なお、翻訳を行う必要が全くない場合には、ホストコンピュータ 3 の翻訳プログラム 3 p y および翻訳エンジン E M 3 などを削除しておけばよい。

## 〔第二の実施形態〕

図 8 は第二の実施形態における端末装置 2 の側に記憶されているプログラムおよびデータの例を示す図、図 9 は第二の実施形態におけるホストコンピュータ 3 の側に記憶されているプログラムおよびデータの例を示す図である。

## 【 0 0 6 1 】

第二の実施形態において、通信システムの全体の構成は第一の実施形態の場合と同様である。しかし、端末装置 2 およびホストコンピュータ 3 に記憶されるプログラムおよびデータと、各処理装置 2 1、3 1 による処理内容とが、第一の実

施形態の場合と異なる。

【0062】

すなわち、第一の実施形態では、顔画像データベースFDB、動作制御プログラム3pd、およびアニメーションエンジンEM2がホストコンピュータ3に設けられているが、第二の実施形態では、図8に示すように、端末装置2に設けられている。したがって、図9に示すように、ホストコンピュータ3には、これら顔画像データベースFDB、動作制御プログラム3pd、およびアニメーションエンジンEM2は設けられていない。

【0063】

図10は第二の実施形態における通信システム1の処理の流れを示すフローチャート、図11は端末装置2の処理の流れを示すフローチャート、図12はホストコンピュータ3の処理の流れを示すフローチャートである。

【0064】

図10に示すように、まず、端末装置2Aと2Bとの間で通信を確立する(#40)。各端末装置2の間において、互いの相手の顔画像データFGDを交換する(#41)。

【0065】

各ユーザは会話を開始するに当たり、翻訳が必要か否かを判断する(#42)。この判断は、第一の実施形態ではホストコンピュータ3によって行われるが、第二の実施形態では端末装置2が行う。例えば、ステップ#41において、各端末装置2は、顔画像データとともに、自分の話す言語の情報を相手に送信する。これに基づいて、相手の各端末装置2は翻訳が必要か否かを判断する。

【0066】

会話が行われると、翻訳が必要な場合には、音声データSNDがホストコンピュータ3に送信される(#43)。ホストコンピュータ3は、それを翻訳し(#44)、翻訳した音声データSNDを相手の端末装置2に送信する(#45)。翻訳が必要でない場合には、音声データSNDは相手の端末装置2に送信される(#46)。

【0067】

端末装置 2 は、受信した音声データ S N D に基づいて、動作制御データ D S D を生成する（＃ 4 7）。受信した音声を出力し、生成した動作制御データ D S D に基づいて顔画像データ F G D を動作させ、生成したアニメーションを表示する。（＃ 4 8）。

【 0 0 6 8 】

図 1 1 に示すように、各端末装置 2 においては、相手のユーザの顔画像データ F G D を受信し（＃ 5 1）、自分の顔画像データ F G D を相手に送信する（＃ 5 2）。

【 0 0 6 9 】

このとき、受信した顔画像データ F G D を顔画像データベース F D B に保存し、再度同じ相手と会話をするときにも利用してもよい。

翻訳の必要がある場合には（＃ 5 3 でイエス）、音声データ S N D をホストコンピュータ 3 に送信する（＃ 5 4）。翻訳が必要でない場合には、音声データ S N D を相手の端末装置 2 に送信する（＃ 5 5）。

【 0 0 7 0 】

相手の端末装置 2 またはホストコンピュータ 3 から音声データ S N D を受信すると（＃ 5 6 でイエス）、動作制御データ D S D を生成し（＃ 5 7）、それらによる音声を出力しアニメーションを表示する（＃ 5 8）。

【 0 0 7 1 】

図 1 2 に示すように、ホストコンピュータ 3 においては、端末装置 2 から音声データ S N D を受信すると（＃ 6 1）、翻訳を行い（＃ 6 2）、翻訳された音声データ S N D 2 を相手側の端末装置 2 に送信する（＃ 6 3）。

【 0 0 7 2 】

第二の実施形態によると、動作制御データ D S D をホストコンピュータ 3 ではなく端末装置 2 で生成するので、動作制御データ D S D の送受信が不要である。これにより、通信量をさらに軽減することができる。

【 0 0 7 3 】

第二の実施形態において、翻訳を行う必要が全くない場合には、ステップ＃ 5 3 の判断を行うことなく、音声データ S N D を常に相手の端末装置 2 に送信す

ばよい。その場合に、ホストコンピュータ 3 を介する必要はない。したがって、この場合には、ホストコンピュータ 3 を用いることなく、簡易なネットワークにおいても通信システムを構築することができる。

【0074】

上に説明したいずれの実施形態においても、会話を開始するに当たり、各端末装置 2 において顔画像データ FGD を予め取得しておき、会話中においては動作制御データ DSD に基づいてアニメーションを生成するので、通信量が軽減し、自然な動きのアニメーションを表示することができる。

【0075】

また、上に述べたように、音声データ SND に対応するテキストデータをユーザに対して出力することにより、アニメーション、音声、文字データのマルチメディアを構成することができる。

【0076】

その他、端末装置 2、ホストコンピュータ 3、または通信システム 1 の全体または各部の構成、回路、処理内容、処理順序、通信の手順などは、本発明の趣旨に沿って適宜変更することができる。

【0077】

【発明の効果】

本発明によると、通信量を軽減させ、かつ、自然に近いスムーズな相手のアニメーションを見ながら会話を行うことが可能となる。

【0078】

請求項 2、4、5 の発明によると、言語の異なる者同士であっても、自然に近いスムーズな相手のアニメーションを見ながら会話を行うことが可能となる。

請求項 5 の発明によると、動作制御データの送受信を行わないので、通信量をさらに軽減することができる。

【0079】

請求項 6 の発明によると、複数の端末装置をネットワークを介して接続することにより、アニメーションを見ながら会話を行うことのできる通信システムを簡単に構成することができる。

【図面の簡単な説明】

【図 1】

本発明に係る第一の実施形態の通信システムの全体の構成を示すブロック図である。

【図 2】

端末装置の側に記憶されているプログラムおよびデータの例を示す図である。

【図 3】

ホストコンピュータの側に記憶されているプログラムおよびデータの例を示す図である。

【図 4】

アバターとなる顔画像データの例を示す図である。

【図 5】

通信システムの処理の流れを示すフローチャートである。

【図 6】

端末装置の処理の流れを示すフローチャートである。

【図 7】

ホストコンピュータの処理の流れを示すフローチャートである。

【図 8】

第二の実施形態における端末装置の側に記憶されているプログラムおよびデータの例を示す図である。

【図 9】

ホストコンピュータの側に記憶されているプログラムおよびデータの例を示す図である。

【図 10】

通信システムの処理の流れを示すフローチャートである。

【図 11】

端末装置の処理の流れを示すフローチャートである。

【図 12】

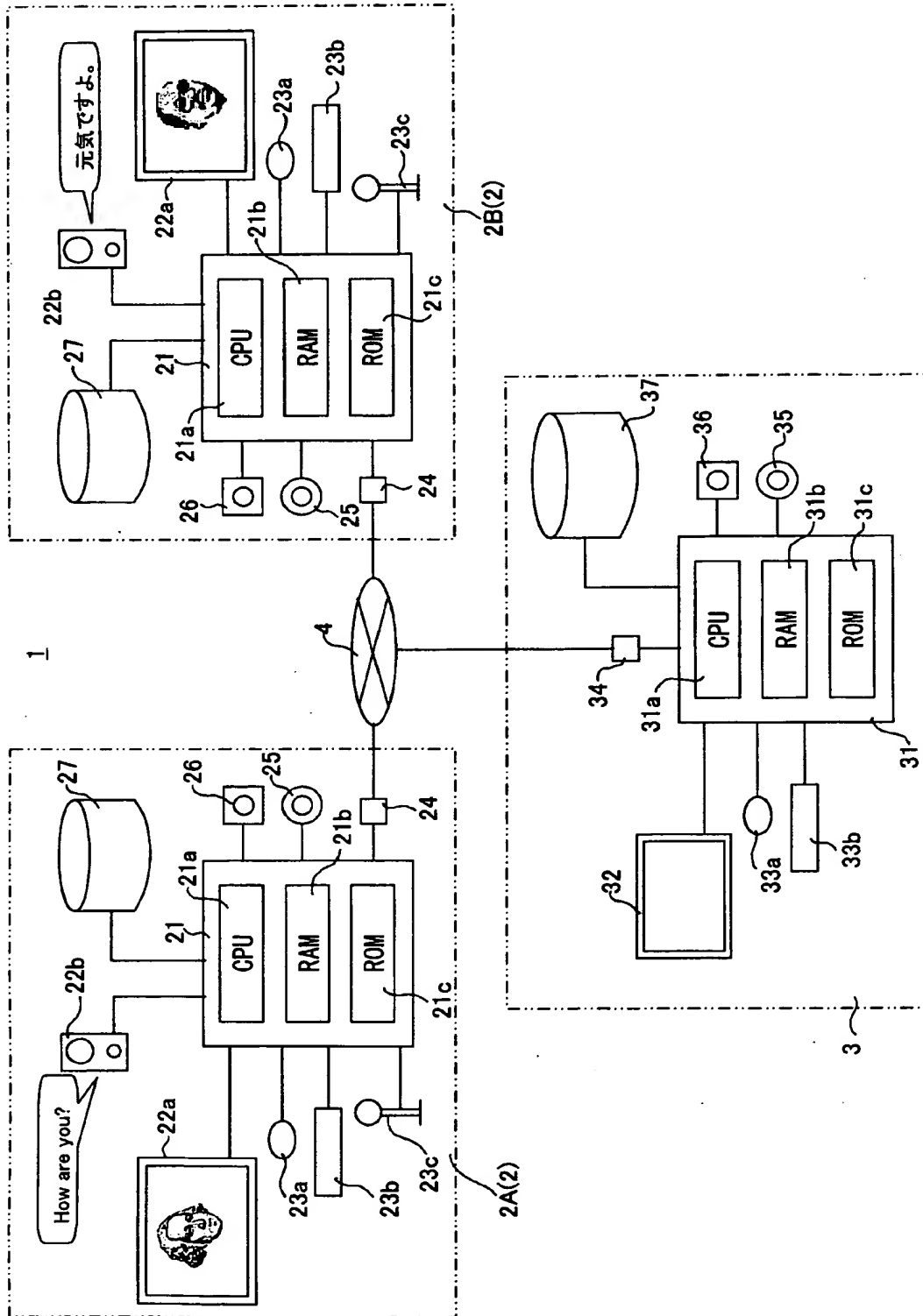
ホストコンピュータの処理の流れを示すフローチャートである。

【符号の説明】

- 1 通信システム
- 2 端末装置
- 3 ホストコンピュータ
- 21 処理装置（音声を送受信する手段、3次元の画像データを受信する手段、動作制御データを受信する手段）
- 22a 表示装置（アニメーションを表示する手段）
- 24 通信制御装置（送受信する手段）
- 31 処理装置（動作制御データを生成する手段、送信する手段）
- 34 通信制御装置（送信する手段）
- 37 磁気ディスク装置（記憶手段）
- EM1 表示制御部（アニメーションを表示する手段）
- EM2 アニメーションエンジン（動作制御データを生成する手段）
- EM3 翻訳エンジン（翻訳する手段）

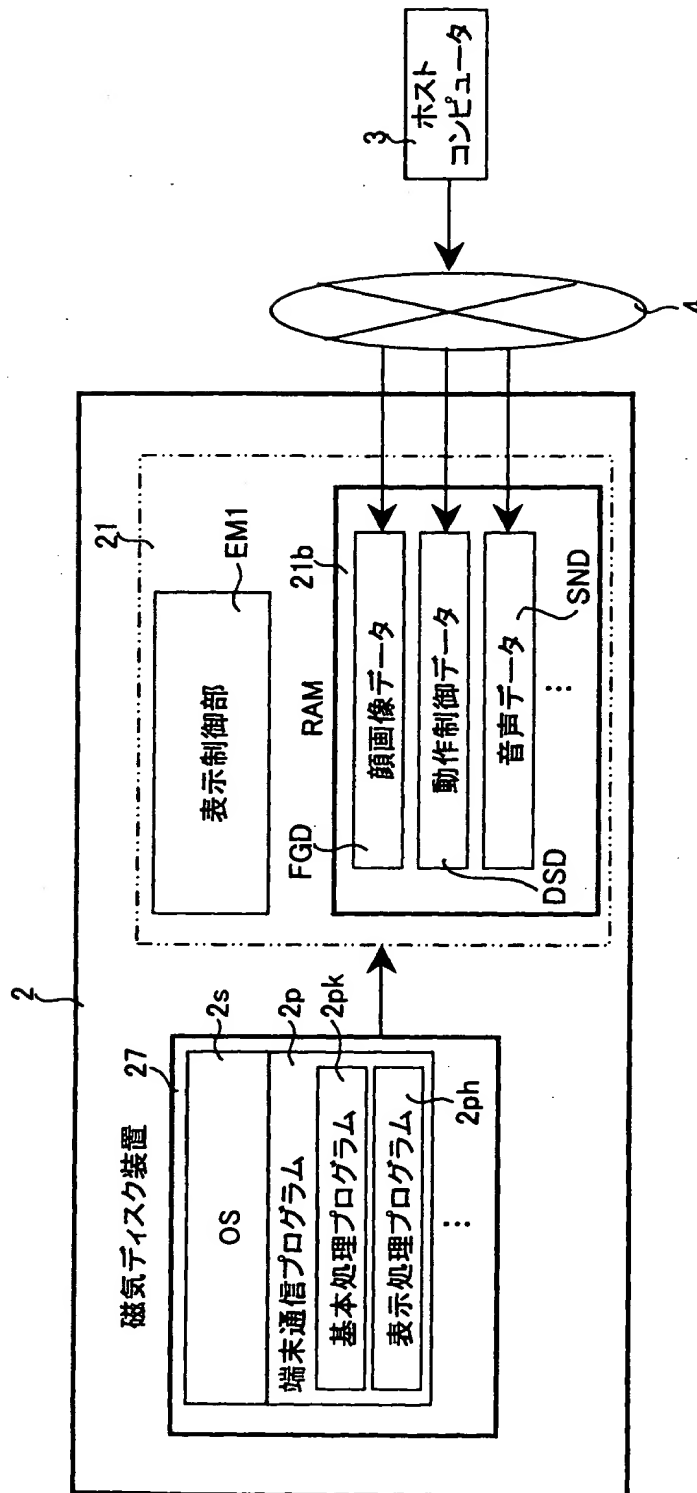
【書類名】 図面

【図 1】

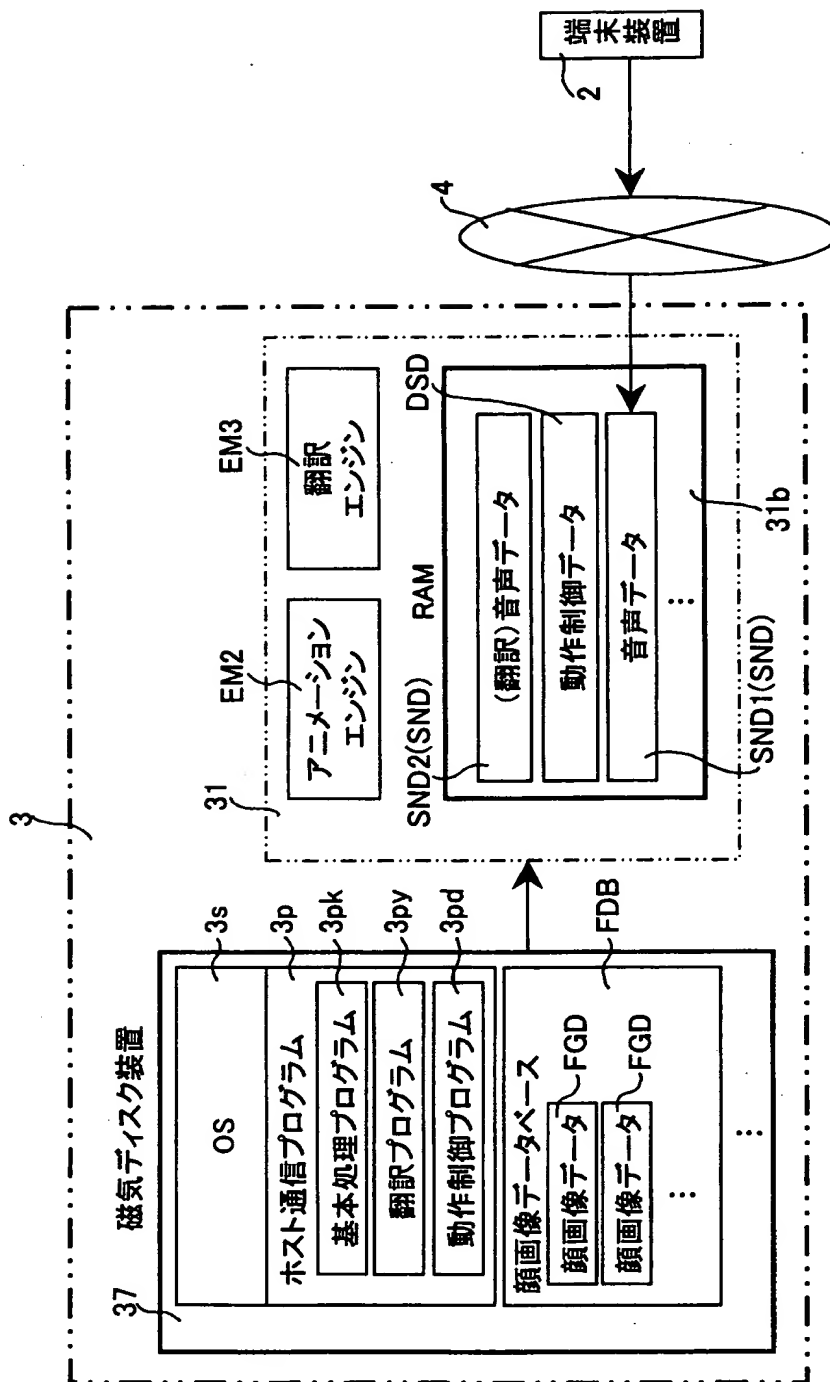




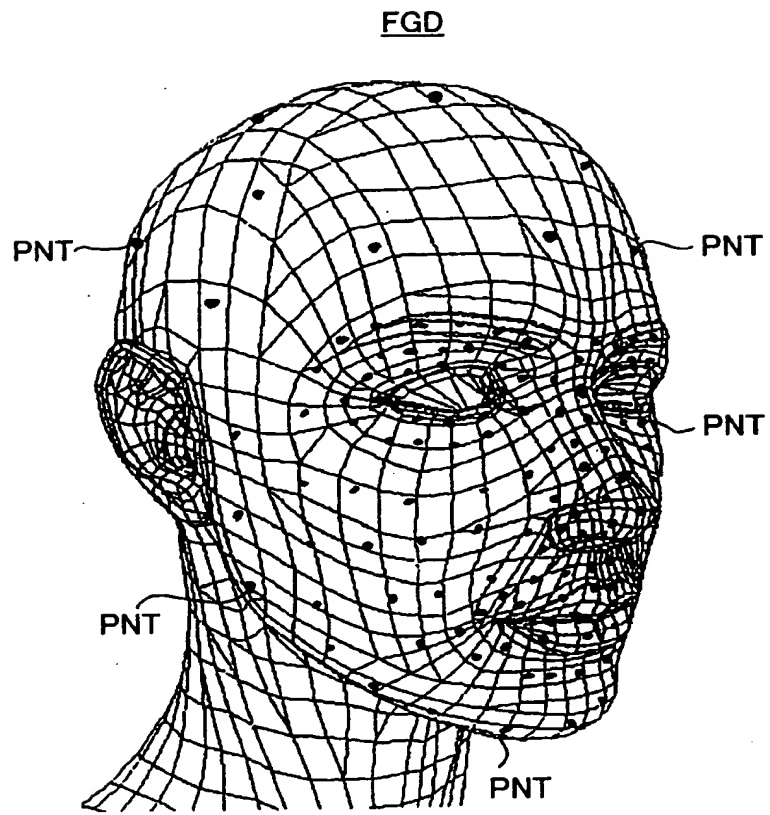
【図 2】



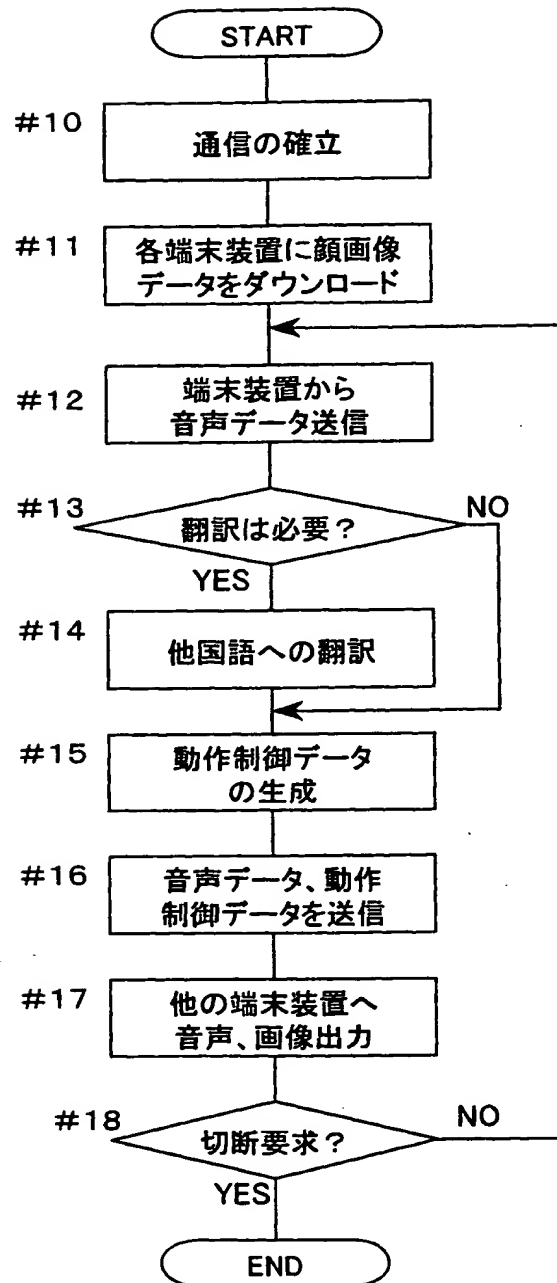
【図 3】



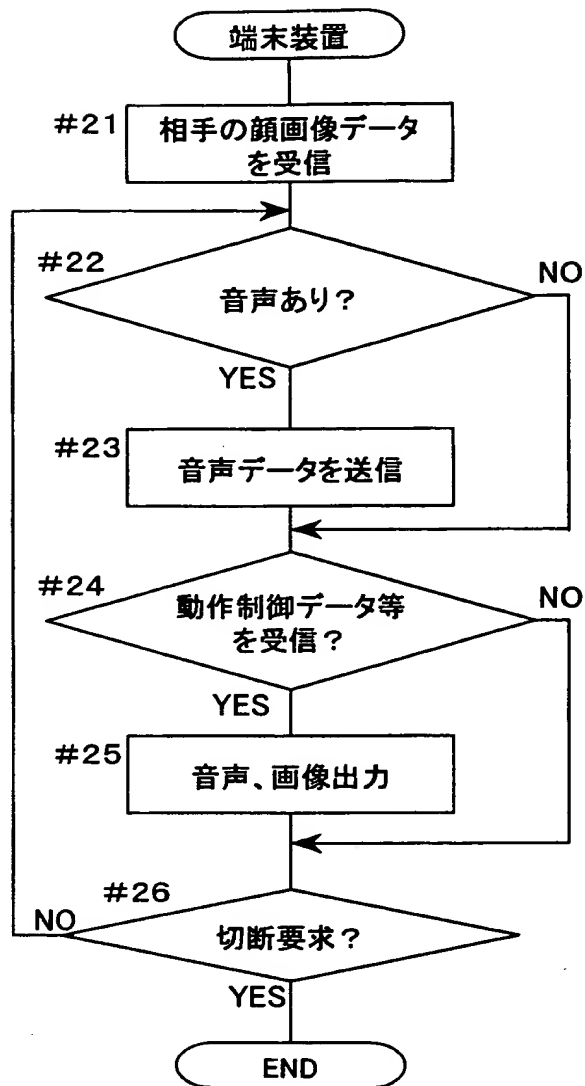
【図4】



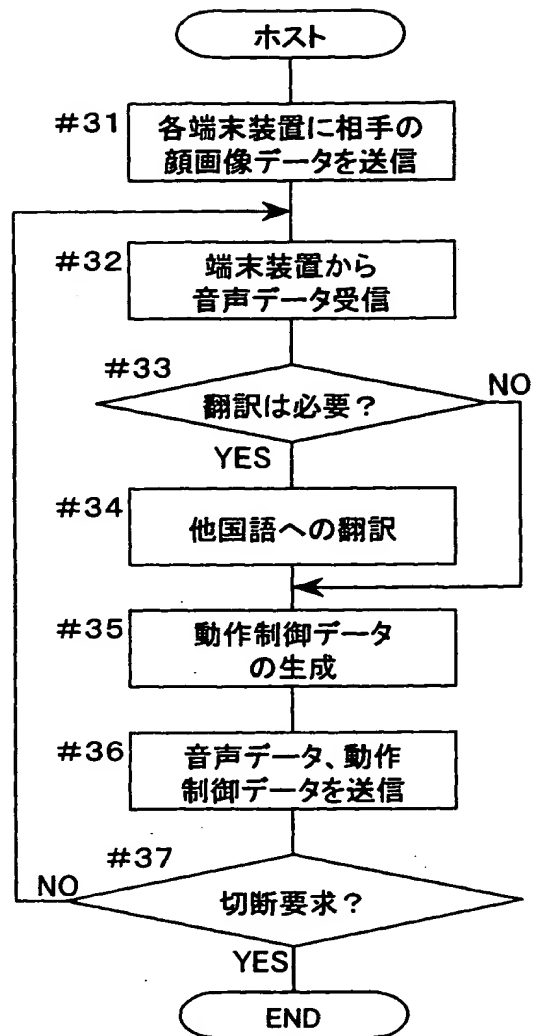
【図 5】



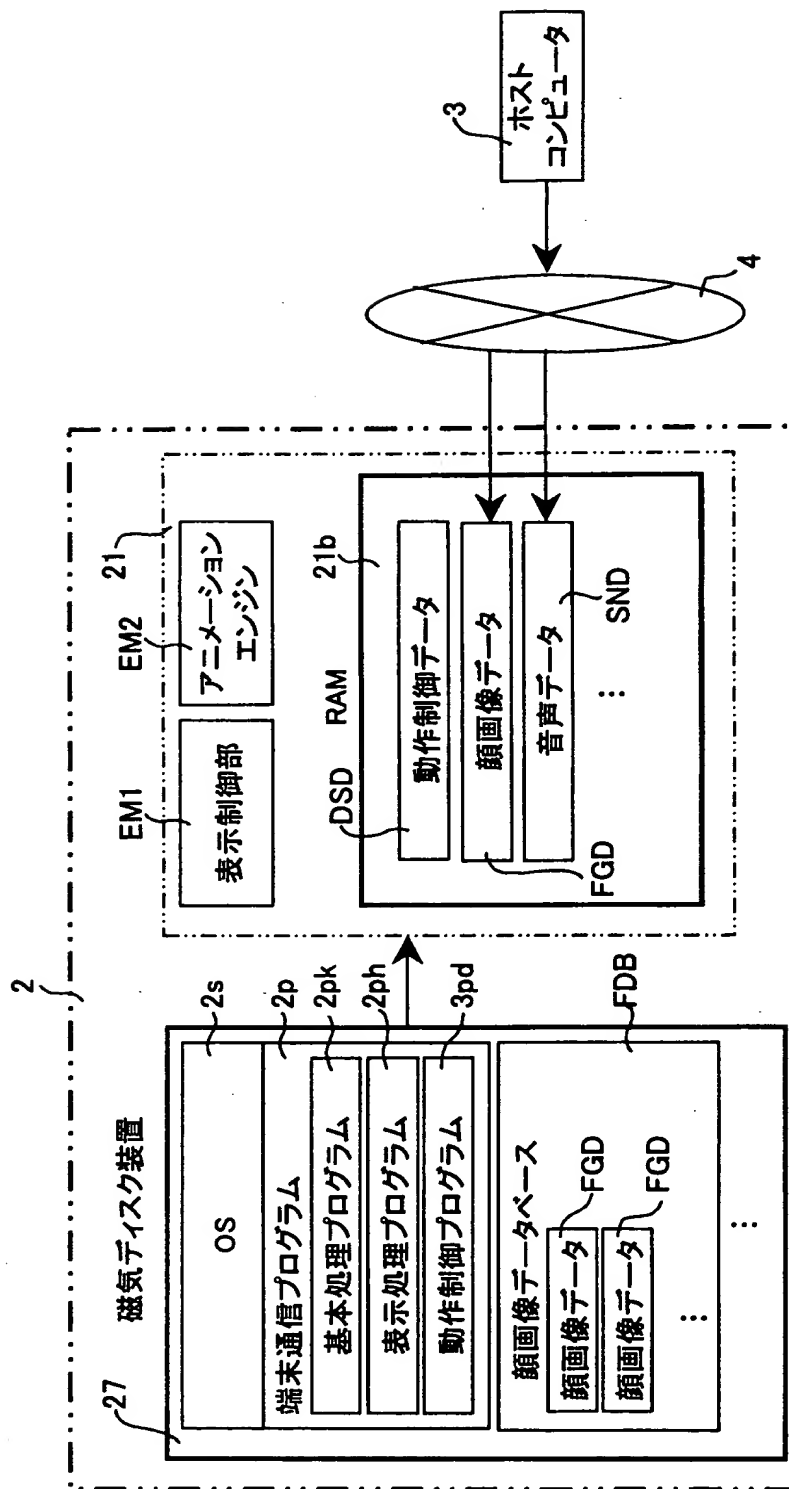
【図 6】



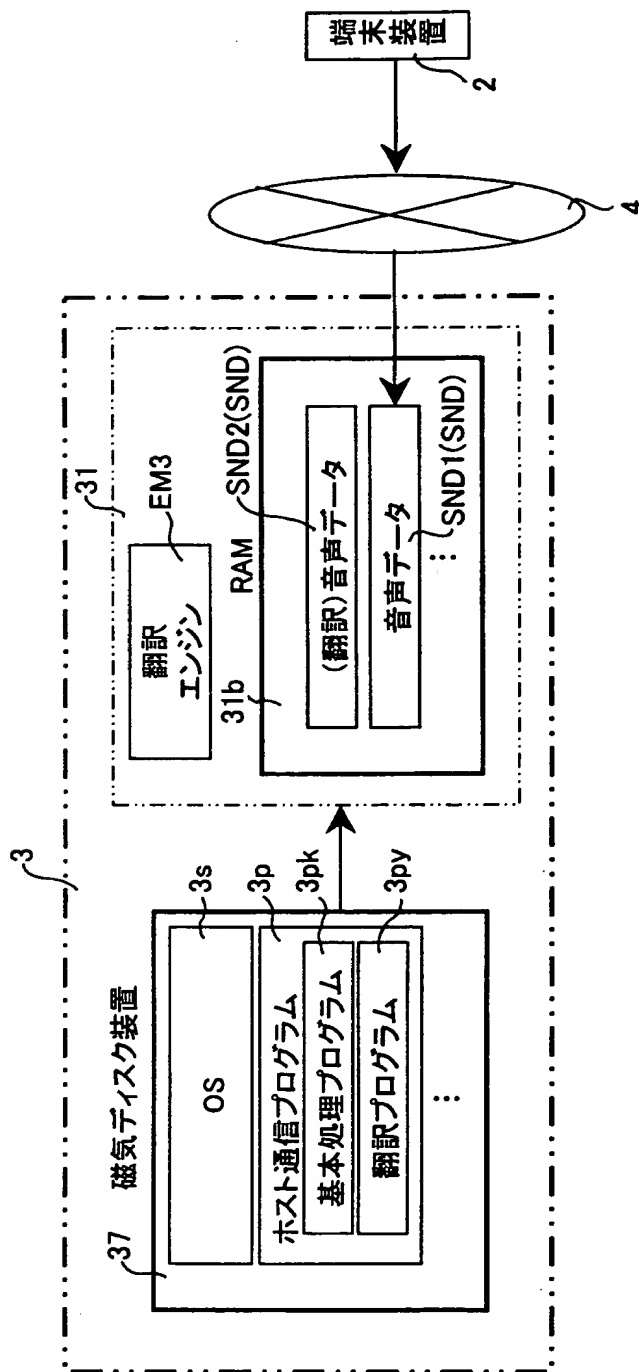
【図 7】



【図 8】

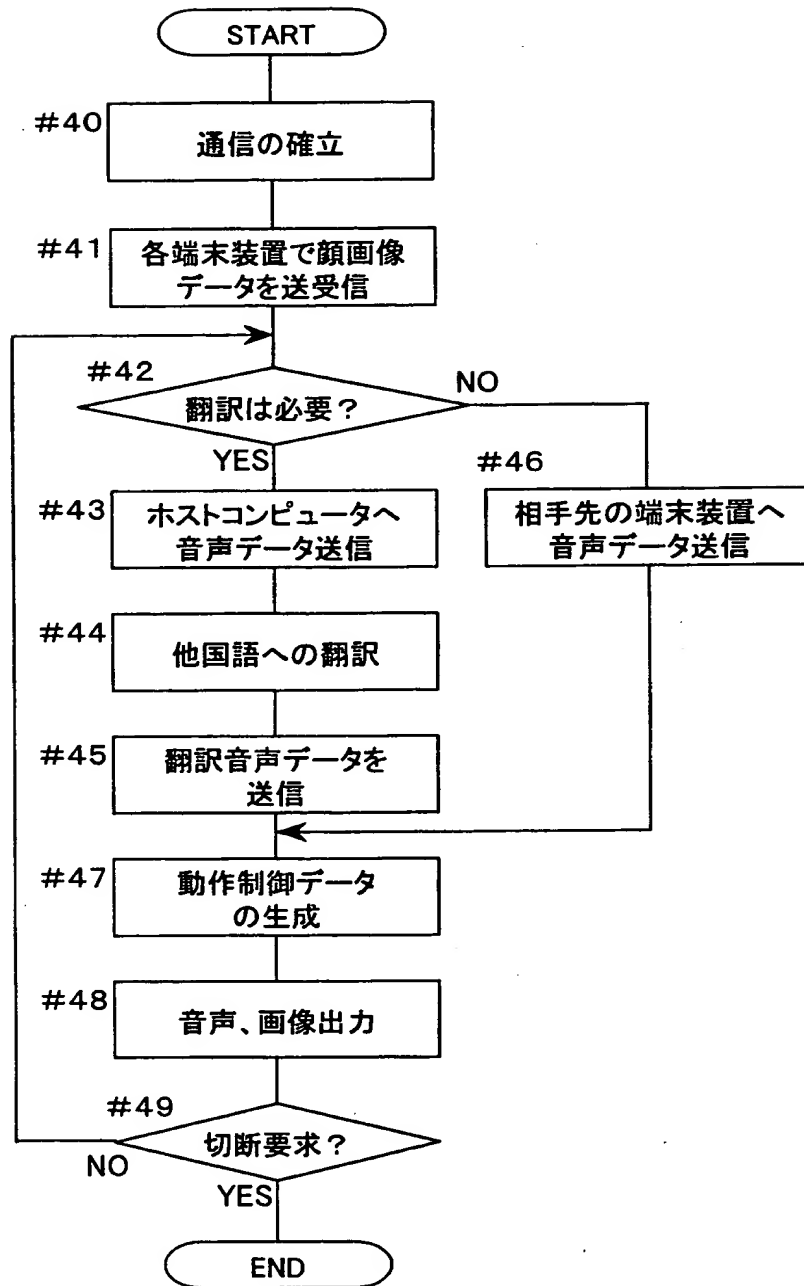


【図9】

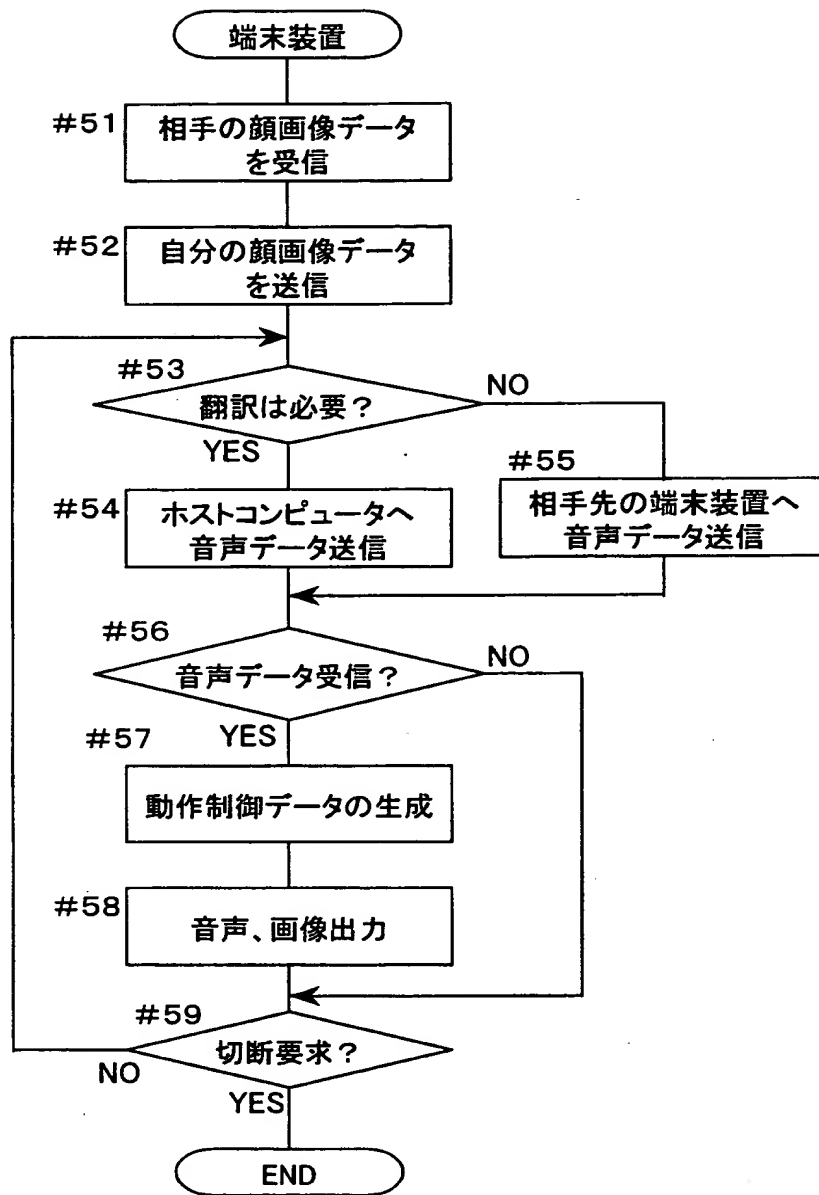




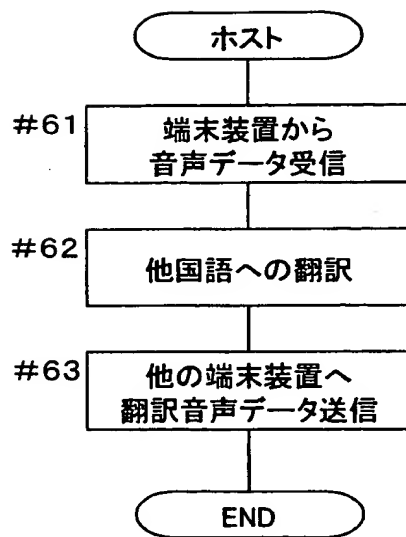
【図 10】



【図 11】



【図 12】



【書類名】 要約書

【要約】

【課題】 通信量を軽減させ、且つ自然に近いスムーズな相手のアニメーションを見ながら会話を行うことを可能にすること。

【解決手段】 ホストコンピュータ 3 と複数の端末装置 2 とを有し、端末装置 2 には、音声を送受信する手段と、3 次元の画像データを受信する手段と、画像データを動作させるための動作制御データを受信する手段と、動作制御データに基づいて画像データを動作させて生成される 3 次元のアニメーションを表示する手段とが設けられ、ホストコンピュータ 3 には、音声を送受信する手段と、受信した音声に基づいて動作制御データを生成する手段と、通信の一方の端末装置に係る画像データおよび動作制御データを通信の他方の端末装置に送信する手段とが設けられる。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000006079]

1. 変更年月日 1994年 7月20日

[変更理由] 名称変更

住 所 大阪府大阪市中央区安土町二丁目3番13号 大阪国際ビル

氏 名 ミノルタ株式会社